

CLAIMS

- 1 1. A method for efficiently organizing and distributing parity blocks among storage de-
 2 vices of an array coupled to a storage system, the method comprising the steps of:
 3 configuring the array with a first predetermined number of devices;
 4 dividing each device into blocks;
 5 organizing the blocks into stripes, wherein each stripe contains a first device that
 6 holds only row parity blocks, a second device that holds only diagonal parity blocks and
 7 remaining devices that hold data blocks, and wherein the number of blocks on each de-
 8 vice of each stripe is equal;
 9 appending an extra parity block within each stripe to the second device holding
 10 the diagonal parity blocks within that stripe, such that an uneven device length arises
 11 within the stripe; and
 12 assigning each data block to one row parity set and one diagonal parity set, such
 13 that any one or any combination of two devices can fail concurrently and data on the
 14 failed devices can be reconstructed.
- 1 2. The method of Claim 1 wherein the step of dividing comprises the step of dividing
 2 each device into blocks of fixed size.
- 1 3. The method of Claim 1 wherein the first predetermined number of devices is n .
- 1 4. The method of Claim 3 wherein n is any prime number plus 2.
- 1 5. The method of Claim 3 further comprising the step of distributing the parity blocks
 2 among the devices of each of n different adjacent stripes to create a super-stripe contain-
 3 ing n stripes that balances the sizes of the devices.
- 1 6. The method of Claim 5 further comprising the step of repeating the super-stripe over
 2 an entire length of the devices.

1 7. The method of Claim 3 wherein the step of organizing comprises the step of organiz-
2 ing the blocks into stripes, each stripe comprising $(n \times (n-3)) + 1$ blocks.

1 8. The method of Claim 1 wherein the step of appending comprises the step of obviating
2 additional exclusive-OR (XOR) operations to thereby reduce the amount of computation
3 needed to encode diagonal parity.

1 9. The method of Claim 8 further comprising the step of computing parity contained in
2 the parity blocks using addition over fixed size multi-bit fields.

1 10. The method of Claim 9 wherein the fixed size of the multi-bit fields is one of 8, 16,
2 32, 64 and 128 bits.

1 11. The method of Claim 1 wherein the step of organizing comprises the steps of:
2 exchanging a first block at a first position on a device to a second position of a
3 second block on the device without affecting contents of the blocks; and
4 altering assignment of the blocks to parity sets.

1 12. A system for efficiently organizing and distributing parity blocks among devices
2 coupled to a storage system, the system comprising:
3 an array having a plurality of devices, each device divided into blocks of fixed
4 size, the blocks organized into stripes, wherein each stripe contains a first device that
5 holds only row parity blocks, a second device that holds only diagonal parity blocks and
6 remaining devices that hold data blocks and wherein the number of blocks on each device
7 of each stripe is equal;

8 a storage operating system including a device storage layer configured to imple-
9 ment a distributed parity technique that appends an extra parity block within each stripe
10 to the second device holding the diagonal parity blocks within that stripe, such that an
11 uneven device length arises within that stripe, the device storage layer assigning each
12 data block to one row parity set and one diagonal parity set, such that any one or any

13 combination of two devices can fail concurrently and data on the failed devices can be
14 reconstructed; and

15 a processing element configured to execute the storage operating system to
16 thereby invoke storage access operations to and from the array in accordance with the
17 distributed parity technique.

1 13. The system of Claim 12 wherein the device storage layer further distributes the parity
2 blocks among the devices of each stripe to create a super-stripe that balances the sizes of
3 the devices.

1 14. The system of Claim 13 wherein the devices are disks and wherein the device storage
2 layer is a RAID layer.

1 15. The system of Claim 14 wherein the RAID layer is implemented in logic circuitry.

1 16. The system of Claim 13 wherein the devices are one of video tape, magnetic tape,
2 optical, DVD and bubble memory devices.

1 17. The system of Claim 13 wherein the devices are media adapted to store information
2 contained within the data and parity blocks.

1 18. The system of Claim 12 wherein the fixed size of the blocks is 4k bytes.

1 19. The system of Claim 12 wherein the fixed size of the blocks is a power of two num-
2 ber of bytes.

1 20. The system of Claim 19 wherein the high power of two bytes is one of 4k, 8k, 16k,
2 32k, 64k, 128k and 256k bytes.

1 21. Apparatus for efficiently organizing and distributing parity blocks among disks of a
2 disk array coupled to a storage system, the apparatus comprising:

3 means for configuring the disk array with a first predetermined number of disks;

4 means for dividing each disk into blocks;

5 means for organizing the blocks into stripes, wherein each stripe contains a first
6 disk that holds only row parity blocks, a second disk that holds only diagonal parity
7 blocks and remaining disks that hold data blocks, and wherein the number of blocks on
8 each disk of each stripe is equal;

9 means for appending an extra parity block within each stripe to the second disk
10 holding the diagonal parity blocks within that stripe, such that an uneven disk length
11 arises within the stripe; and

12 means for assigning each data block to one row parity set and one diagonal parity
13 set, such that any one or any combination of two disks can fail concurrently and data on
14 the failed disks can be reconstructed.

1 22. The apparatus of Claim 21 further comprising means for distributing the parity
2 blocks among the disks of each stripe to create a super-stripe that balances the sizes of the
3 disks.

1 23. The apparatus of Claim 22 further comprising means for repeating the super-stripe
2 over an entire length of the disks.

1 24. The apparatus of Claim 21 wherein the means for appending comprises means for
2 obviating an additional exclusive-OR operation to thereby reduce the amount of compu-
3 tation needed to encode diagonal parity.

1 25. A computer readable medium containing executable program instructions for effi-
2 ciently organizing and distributing parity blocks among disks of a disk array coupled to a
3 storage system, the executable program instructions comprising program instructions for:
4 configuring the disk array with a first predetermined number of disks;

5 dividing each disk into blocks;
 6 organizing the blocks into stripes, wherein each stripe contains a first disk that
 7 holds only row parity blocks, a second disk that holds only diagonal parity blocks and
 8 remaining disks that hold data blocks, and wherein the number of blocks on each disk of
 9 each stripe is equal;
 10 appending an extra parity block within each stripe to the second disk holding the
 11 diagonal parity blocks within that stripe, such that an uneven disk length arises within the
 12 stripe; and
 13 assigning each data block to one row parity set and one diagonal parity set, such
 14 that any one or any combination of two disks can fail concurrently and data on the failed
 15 disks can be reconstructed.

1 26. The computer readable medium of Claim 25 further comprising program instructions
 2 for distributing the parity blocks among the disks of each stripe to create a super-stripe
 3 that balances the sizes of the disks.

1 27. The computer readable medium of Claim 26 further comprising program instructions
 2 for repeating the super-stripe over an entire length of the disks.

1 28. The computer readable medium of Claim 25 wherein the program instruction for ap-
 2 pending comprises program instructions for obviating additional exclusive-OR operations
 3 to thereby reduce the amount of computation needed to encode diagonal parity.

1 29. A method for efficiently organizing and distributing parity blocks among storage de-
 2 vices of an array coupled to a storage system, the method comprising the steps of:
 3 configuring the array with n storage devices;
 4 dividing each storage device into blocks;
 5 organizing the blocks into stripes, wherein each stripe comprises $(n \times (n-3)) + 1$
 6 blocks, each storage device containing a pattern of data and parity blocks within a stripe,

7 the parity blocks being one of row and diagonal parity, the storage device containing the
8 diagonal parity is different in each pattern recurrence.

1 30. The method of Claim 29 where n is any prime number plus 2.